

統計調査における地点情報の把握による統計 の情報価値の新たな展開可能性について

森, 博美 / MORI, Hiromi

(出版者 / Publisher)

法政大学経済学部学会

(雑誌名 / Journal or Publication Title)

The Hosei University Economic Review / 経済志林

(巻 / Volume)

78

(号 / Number)

3

(開始ページ / Start Page)

249

(終了ページ / End Page)

281

(発行年 / Year)

2011-02-25

(URL)

<https://doi.org/10.15002/00007104>

統計調査における地点情報の 把握による統計の情報価値の 新たな展開可能性について

森 博 美

要 旨

近年の精度向上と情報端末の低廉化に伴い、全地球測位システムGPS（Global Positioning System）情報は、現在、多方面に活用されている。このような技術的展開を受けて、本稿では、GPSコードの統計への活用による調査票情報が潜在的に保有する情報価値の開拓並びに拡張に関する方法的考察を行なう。

近代統計の黎明期に実施されていた表式調査と異なり、調査対象単位を調査票によって把握するいわゆる個票調査においては、個々の調査票情報（統計原単位情報）は、調査区コードという明示的な形で、調査対象単位が持つ位置情報を保有している。しかし、現実には、調査区で把握された世帯、事業所、それに企業といった調査対象単位が持つ場所情報は、調査票情報においては、それらが本来的に有するそれぞれの地点情報としてではなく、調査区コードによって集約的に表記されているに過ぎない。

本論文では、調査対象をこれまでのように集団としてではなくあくまでも個体から構成されるものとして把握するという視点から、個票調査の特性並びにそれが調査個票情報に付与する固有の情報特性に注目する。そし

て本稿では、調査対象を構成する各单位が本来的に備えている地点情報を、調査区コードという面的広がりを持った情報としてではなく、個々の住戸や事業所等の所在地情報から得られるGPSコードとして捉えることを提唱している。このような視点に立つことによってはじめて、これまでの調査区コードによる位置情報把握に起因する様々な統計の利用上の制約が明らかにできるとともに、統計の作成面、利用面での新たな統計の展開が可能となる。

【キーワード】 GPS, 調査票情報, 調査区コード, パネルデータ, 表式調査, 産業集積

まえがき

個人や世帯、それに企業や組織といった社会を構成する各主体の属性、活動、活動結果、さらには行動やそれを支配する意識は、時間や空間を貫いて相互に連関を持ちながら、一つの時空間を作り出している。統計が究極的に反映すべき対象という意味でこれを筆者は統計把握空間と呼んでいる〔森（7）p.16〕。統計把握空間においては、時間ならびに空間に関する情報は、象徴的な意味で、社会の構成要素である統計対象単位の個体としての存在の場を規定するいわば座標軸を与えるものである。

このうち時間情報については、静態統計における把握時点、動態統計における参照期間（reference period）、反復的横断面データ（repeated cross-sectional data）によって編成される時系列統計の時点情報、さらには縦断面データ（longitudinal data）におけるそれぞれの参照時点といった形で、調査対象単位に関する統計的把握の諸側面を規定する変数としてとりあげられてきた。その一方で、調査対象単位の存在の場を規定する情報要素である位置情報については、都道府県や市町村といった地域別集計における位階的地域表章レベルとして取り上げられてきただけである。

本稿ではまず、統計調査が調査対象単位が現実存在する場所に関する

情報をこれまでどう捉え、また捉えてこなかったかを、表式調査、個票調査といったこれまでの統計の作成方式と関連づけながら論じる。さらに、それらの検討結果並びに近年の情報技術の展開をふまえつつ、調査票情報が本来潜在的に有する情報量の開拓、拡張の可能性について考察してみることにはしたい。

1. 統計調査と個体情報

(1) 表式調査における場所情報

1) 黎明期における統計的把握－表式調査－

明治4（1871）年に制定された戸籍法に基づいて実施された全国一斉人口検査（いわゆる「戸口調査」）や太政官布告による「府県物産表」など明治初期の統計調査の多くは、表式調査として実施された。この調査方式は調査対象単位を直接個別に把握するのではなく、所定の事項に関して、集計表の様式で予め作成された表式の各枠目（セル）を、集落等における記入担当者が既存の資料からの転記あるいは独自に把握した集計値によって埋め、それらの計数が市町村、市郡、さらには府県、そして最終的には全国といった行政の組織系統を通じて積み上げられることで統計に編成されるものである。それは、「観察者が出来上った表式を手にして表式中に予定せる類別に該当せるものがある場合には、この該当せる事実につき、その相当欄へ差当り印をつけ、次にその印を集計して求める数値を発見するといふ方法で観察結果を決定するもの」〔(10) 98頁、大橋訳235-6頁〕であり、「調査票の過程を省略し、一定の地域を単位として直接集計表形式への記入を要求する調査法」〔(5) 122頁〕である。

このような表式調査として実施された「戸口調査」について鮫島龍行は、「その調査方法は江戸時代の人別改めの方式を出ることなく、きわめて前近代的なもの」〔(3) 31頁〕、「集計組織の未熟な初期の発展段階ではむしろ必

然的な情報収集の方式」〔(3) 27頁〕,さらには「歴史的には,表式調査は集計組織の整備されていない初期段階に適応した方式である」〔(3) 343頁〕と特徴づけている。

ところで,表式調査によって作成された統計は,次のような問題点を内在させていた。特に初期の調査においては,静態調査事項についても調査時点が調査実施者からは明示的に指示されていなかった。また戸口調査の場合,それが兵役への動員という行政目的も同時に持っていたことから,男と女とで年齢階級区分が異なっていた。一方,「物産表」については,その所管官庁である民部省は,明治3年の通達第623号において,単に「府県管内物産取調」は,「山川海陸ノ物産並人工ニ係り候品類トモ其管内取調不洩様記載可致事」として,米,雑穀,塩類など29の品目について「一歳ノ総額」を府県単位で調査し,「取調日数三十日を限り可差出事」を命ずるというように,単に把握すべき物産の品目名と報告期限を指示していただけてであった。そこでは,調査する物産品目の範囲や定義が必ずしも統一されておらず,それらの具体的な設定は各府県の担当者あるいは末端の記入者に委ねられていた。また,同通達中の「総額」についても,物的単位あるいは貨幣単位のいずれによる総額とするかの具体的な指示もなされていなかった。このような統計把握上の問題を内在させた表式調査について,鯨島はそれを,「統計表の形式さえととのっていれば,その数字がどんな手続きでえられたかを問わないし,また問うことのできない方式」〔(3) 27頁〕として特徴づけている。

「物産表」については,現物単位あるいは貨幣単位という総額の測定単位の不統一,さらには調査対象品目の把握対象範囲の不統一から,府県によって報告様式が異なり,品目によっては全国の総生産額の算定さえも困難である場合も散見された。「原理的にいえば,府県統計は市町村表式の各数値を各項目ごとに合算したものにすぎない。全国統計は,さらに府県統計の表式を各項目ごとに合算したものにすぎない。そこで表式調査方式のもとでは,全国計ができる以前に府県統計・市町村統計はできあがって」お

り、従って、「明治期の主要中央統計である『農商務統計表』の記載事項よりも、地方統計である『府県統計書』の記載事項のほうが情報が豊富」〔(3) 328頁〕であったのである。

そこでは、「報告のあったものだけを合算し製表するというのが表式調査における支配的な思想」〔(3) 76頁〕であった。このため、「調査項目を各種の目的に応じて自由に製表することはできないし、調査上の誤りを修正することもできない」〔(3) 343頁〕。なお森は、表式調査の調査特性について、対象反映性との関連で「表式調査の場合、統計原単位情報がもともと集計量であることから、調査項目間の関連性はその中に完全に埋没している。従って表式調査は、統計把握空間に関する認識資料としての利用可能性面で大きな制約を持ち、たとえ調査が正確に実施されたとしても、統計把握空間の若干の側面についての単純集計結果の断片的羅列以上の情報を提供するものではない」〔(7) 16頁〕としている。

2) 表式調査における場所情報

このような調査特性を持つ表式調査は、場所情報をどのように捉えていたのだろうか。

「物産表」については、明治16(1883)年に制定される「農商務通信規則」によって、初めて「通信事項」に関する全国統一様式が定められる。報告様式が統一されたことで、各府県から報告される計数の全国総計への積上げ過程においてかつて存在していた障碍はようやく取り除かれた。なお、通信規則の制定から次世代型の調査方式である個票調査が部分的に導入される明治27(1894)年までの約10年間に、わが国の統計史における表式調査の完成期として捉えている〔(3) 54頁〕。

表式調査の場合、把握された計数は行政組織の系統を通じて積み上げられ、最終的に全国総計が得られる。末端の作成現場で作成される報告様式には、場所情報として作成市町村等の名称が記載されるだけで、それは、『府県統計書』等の府県総括統計においてはじめて域内の地域に関する情報として意味を持つことになる。同様にその上位レベルの府県という地域単

位は、『農商務統計』等の全国レベルでの総括統計における府県別内訳として意味を持つ。そこでは、それぞれの内訳構成そのものには特に重要な意味は与えられておらず、あくまでも全国総計を得るための通過的手続きの中で各地域レベルでの表章が行われているに過ぎない。表式調査が支配的な調査形態であった当時の統計調査の企画・実施者にとっては、徴税あるいは国力の把握といった行政遂行上必要な情報の獲得こそが主たる関心事であった。そこで意味を持つのはあくまでも全国総計あるいは特定の年齢等の属性別内訳であり、地域別表章は、レベルを異にする地域データの積上げという表式調査の統計作成過程から派生するいわば「副産物」的性格のものであった。そこでは、地域比較といった固有の分析目的のために特別の意図を持って地域情報が把握されたわけでは必ずしもない。

（２）個票調査における場所情報

１）表式調査から個票調査へ

蘭学を通じて海外の統計事情に精通していた明治期初期の指導的統計家であった杉亨二や呉文聡は、表式調査の限界を早くから見抜いていた。そして彼らは、政府統計への個票調査の導入に積極的な役割を果たした。

明治４（1871）年に太政官正院に設置された製表課の責任者として迎えられた杉は、戸口調査という戸籍に基づく表式調査による人口統計の作成には当初から批判的であった。そして明治12（1879）年12月31日午後12時現在で実施されたのが、「甲斐国現在人別調」である。

この調査は、杉自身の意図するところでは、個票方式による全国人口調査の試験調査としての性格を持っていた。それは、人口統計分野における個票調査の嚆矢ともいべきものであった。なお、わが国の場合、諸般の事情により、近代的な人口センサスの実施は、大正９（1920）年まで待たねばならなかった〔(4) 171頁, (6) 5頁〕。

一方、産業統計の分野では、呉の尽力もあり、明治27（1894）年の改革

で「工場票」と「会社票」という工場と会社をそれぞれ単記する「小票」が産業調査に導入された。こういった小票の採用を鮫島は、「明治期を通じて統計情報の支配的な収集形式であった表式調査から近代的な調査票形式、いわば点計調査形式への最初の移行を意味していた」〔(3) 59頁〕と高く評価している。

個票調査が部分的に導入された結果、明治末期の産業統計では、表式調査による農商務統計と、明治42（1909）年以降「工場統計報告規則」に基づき個票調査として実施された産業統計という二つの性格を異にする調査方式による統計調査が並存することになった。

わが国の政府統計の作成システムが全面的に個票調査という基盤の上に確立するようになるのは、大正期になってからである。すなわち、「集団とそれを構成する単位」の概念が官庁統計の中に認識されてきたのは、ようやく大正期にはいつてからのことである。つまり、日本の統計思想の一般水準は、大正期にはいつて、ようやく点計調査の概念、すなわち集団を構成する単位を点計することによって集団の量、特性を知るという認識段階に到達した」〔(3) 279頁〕。

個票調査はそれまでの表式調査と異なり、調査個票情報として個体ベースでの統計原単位情報の獲得を可能にする。そのようにして、得られた調査結果は、調査項目間の構造解明を含め、多面的な分析を可能にする情報面での潜在性を持つ。もっとも、情報処理能力の制約から、個票調査によって把握された母集団の構成要素である個体に関する調査票情報が持つ潜在的価値が現実のものとして全面的に開花するのは、情報処理技術が飛躍的に発展する比較的最近になってからのことである。

2. 諸家における「場所」概念

本節では、社会統計学における代表的な理論家としてその後の統計学の展開に大きな影響を与えたマイヤー（Georg von Mayr）と蜷川虎三の所説

をとりあげ、彼らが統計調査における場所概念をどう捉えていたかを概観する。

(1) マイヤーにおける場所規定

バイエルン統計局長であり後にミュンヘン大学の統計学と財政学の正教授を勤めたマイヤー（1841～1925）は、ドイツにおける政府統計の指導的統計家であり、また当時のドイツ社会統計学派の中心的理論家であった。当時の統計調査の主要な形態は、いうまでもなく調査対象単位を悉皆的に把握するセンサス型の調査であった。

そのような中で彼は、統計調査における場所規定の意義について、(a)「具体的な調査の場所」を規定すること、(b)「調査機関が活動の手をさし拡げ、観察すべき集団の要素をその範囲内で捉へねばならぬ調査区域の規定と限定にとって重要」であり、「一次統計にあつては、…調査機関の観察分野の場所限定によって、観察すべき場所区域が特別に構成されねばならない（調査区）」〔(9) 85頁, 大橋訳205-6頁〕と指摘している。

さらにマイヤーは、実査で用いる調査票に場所を示す情報（地点コード）を記載しておくことの意義について、(c) 観察結果を記入すべき場所、すなわち実査の場所とその範囲を明確にすること、さらには、(d) 観察結果として得られた個々の調査票について「随時、事実との符合を検証し且つ調査用紙を整然と配列整頓することを可能ならしめるため」〔(9) 98頁, 大橋訳236頁〕に必要である、としている。

これらの指摘から、マイヤーにおける場所概念について、次のように考えることができる。まず (a) は、調査の企画実施者が具体的にどの地域で調査を実施するかという情報を調査従事者に指示するものであり、また (b) と (c) は、場所情報を調査技術上の問題として捉え、調査区設定の必要性を指摘したものである。すなわち、調査実施者の立場から統計論を展開したマイヤーにおいては、重複も把握漏れもなく正確に1回だけ調査対象単位を把握する必要上、調査の全対象範囲を排反的な調査区群に区分す

ることが何よりも必要とされたのである。さらに（d）では、場所情報としての調査区コードは、観察結果の点検、さらにはその後の集計過程での使用からその必要性が説かれている。

このように、マイヤーの場合、場所は把握の対象となる世帯や事業所等の調査対象単位の空間的存在に関する情報そのものの取得というよりはむしろ、正確な調査対象単位の把握並びに調査結果の点検、さらには統計作成の後半過程である集計処理をより効率的に遂行するための純粋に統計作成技術上の操作的概念として位置づけられているように思われる。このような調査技術の発展段階においては、調査対象単位の存在を調査区という個々の調査員が実査可能な範囲の区画として捉えておきさえすれば、調査技術上の要件は充足できたのである。

（2）蜷川の大量・大量観察の四要素における場所規定

それでは、わが国における社会統計学の創始者として知られる蜷川虎三（1897～1981）は、どのような統計観に立ち、その中で場所規定は一体どのような意味を持っていたのであろうか。

1）蜷川における統計観

蜷川は、「其の存在が社会的に規定せられたる集団」をいわゆる「大量」と規定し、「之（大量－引用者）を数量的に認識把握することを大量観察と謂」い、「その結果たる一団の数字」を統計と呼んでいる〔（1）17頁〕。それは、「一定の大量の大きさとその集団性の方向及強度とを一個の社会的事実として、物語る所の数字」〔（1）11頁〕であり、「社会的集団を語る数字」〔（2）31頁〕である。大量という基底概念から演繹的に統計理論を構築した蜷川は、統計の作成過程を純粋に統計技術的に理論化したマイヤーの統計論とこの点において異なる。

2）大量・大量観察の四要素

大量を「社会的に其の存在の規定せられたる集団」とする蜷川にとって、大量は「集団なるが故に個別的存在を其の構成因子」〔（1）98頁〕としてい

る。この構成因子が「単位」なのである。蜷川にとっての関心事はあくまでも大量が持つ「集団としての性質即ち集団性」〔(1) 98頁〕である。

大量が持つ多様な集団特性を明らかにするには、それを構成する一つひとつの特性要素（標識）を集団性の方向として、その部分大量の大きさを集団性の「強度」として計測する〔(1) 98頁〕必要がある。しかし、蜷川にとって単位はそれ自体として意味を持っているのではなく、あくまでも「大量の大きさ」を測る際の単に数え上げられるべき個々の対象に他ならない。

3) 蜷川における場所規定

存在の場所は、存在の時と並んで大量及び大量観察の四要素を構成する。しかし、場所と時という2つの要素は、同じく四要素を構成する単位や標識と異なり、直接的にその社会的存在としての大量がまとう「場」的要素であるといえる。蜷川は言う、「大量は社会的存在である。ゆゑに一定の存在の時と存在の場所を有つことは云うまでもない。大量の単位、標識、存在の時及び存在の場所は、大量に就いて不可欠の要素である」〔(1) 99頁〕と。蜷川は時と場所という2つの要素の必要性の根拠について、単に「社会的存在」という大量の存在のあり方以上の言及は行っていない。

蜷川には、場所規定について、しかるべき社会科学の理論に基づく大量へのアプローチという視覚を窺わせる次のような記述がある。すなわち、彼は、「場所と云っても必ずしも単に地理的或は行政的区画を意味するのではなく、大量が存在する社会的領域をいふのである。普通に地理的或は行政的区画がとられるのは大量観察に於ける技術上の理由からで、必ずしもそれが客観的な大量の存在の場所を示すものではない」〔(2) 38頁〕と述べ、場所について、「必ずしも行政的区画を意味するものではなく、要するに、単位の一団が大量として限定せらるる空間的存在条件たればよい。従って、生産米を市場により、或は又、交通路により、平野によって大量として限定することも、かく規定せざるを得ざる必然性を有つならば、その方が寧ろ正しいのである」〔(1) 160頁〕として、いわゆる部分大量の表章

については、通例、地域表章において採用されている行政区分とは異なる地域表章の可能性あるいはむしろその合理性を示唆している。

蜷川においては、調査の実施範囲さらには結果の表章を行う地域区分はあくまでも依拠する社会科学の理論が自ずと設定するものであり、社会科学の理論に従って「大量の大いさを問題にする限り、大量の存在を時と場所によって限定し、単位を数へ上げる範囲を定めなければならぬ」〔(1) 148頁〕ということになる。

一定時点、一定地域における静態量としての大量の集団性の方向並びに強度の把握こそが統計の目的であるとする蜷川においては、場所とはあくまでも大量の大いさを確定するのに必要な限りでの存在の場を規定するものに他ならない。そこでは、マイヤーに見られる調査技術上の調査区設定の意義、すなわち、調査の対象範囲を排反的な調査区の集合体として設定することで調査対象単位を漏れなく正確に1回限り把握するという調査の技術論理は背景に退いている。要するに蜷川においては、調査実施のための場所の地域区分は、何よりも大量をどう捉えるかに依存するものであり、それは依拠する社会科学の理論によって自ずと決定されるものである。彼には、事後的に行われるであろう多様な分析目的に対応しうような場所情報の持ち方を統計作成技術としてどう構築するかという発想は存在しない。そこでは、単に調査区を地域集団の最小単位とすることだけで十分であり、それらを積み上げることで部分大量を表章することさえできれば、統計としての使命は果たされるのである。

3. 統計における場所情報表章の展開

近代統計調査では、基本的に調査個票によって個体に係る統計原単位情報の収集が行われる。そこでは、社会の中で様々な関係を相互に取り結び、判断、行動する個体のある側面が調査票における調査事項として写し取られる。それは、統計対象単位に関する統計原単位情報として、集計量とし

て編成される統計の出発点となる。

これまでの個票調査では、場所情報は、何よりもまず調査区情報、すなわち一定の地域的広がりを持つ空間情報として調査票情報の一つの変数を形成している。そこでは、同一調査区の調査対象を構成する複数の単位が、その情報を共有している。その後、集計機械の処理能力の向上に伴い、作成公表される結果表が拡充され、特に国勢調査や企業・事業所統計といったセンサスについては、従来からの地域表章に加えて次のような様々な小地域集計が行われるようになる。

（１）小地域集計

わが国では「町村合併法」（昭和28年）や「新市町村建設促進法」（昭和31年）の施行に伴う市町村域の広域化、高度成長に伴う人口の都市集中と地域間格差の顕在化、さらには広域市町村圏での社会基盤整備計画の策定といった政策ニーズなどが、統計についての新たな種類の地域表章を要請することになった。このような中、集計結果に実質的な意味を持たせるため、国勢調査については、昭和35年調査から人口集中地区（DID）についての集計結果が作成、公表されることになる。

人口集中地区による地域別表章の他にも、その後、国勢調査については、様々な小地域集計が導入される。例えば、平成2年からは基本単位区別集計が、また平成7年からは、基本単位区を元に集計した町丁・字等別集計が作成提供されている。

また企業・事業所統計調査についても、平成3年には、調査票に記載された企業・事業所の所在地情報を元に集計した町丁・字別集計が、また平成8年からは調査区別集計も行われるようになった。

（２）地域メッシュ統計

メッシュ統計とは、地域をメッシュといわれる方形（グリッド）に区分し、個々のグリッド内に所在する調査対象単位に関する統計情報を、総計

表1 地域メッシュの区画体系と基準地域メッシュ

区画の呼称	一辺の長さ	区画数	区画コード	備考
第一次地域区画	約80km	176	4 桁	
第二次地域区画	約10km	4,885	4 + 2 桁	第一次地域区画の1辺を8等分
第三次地域区画 基準地域メッシュ	約1 km	386,877	4 + 2 + 2 桁	第二次地域区画の1辺を10等分

や比率といった指標（集計量）として表示する表章方式である。政府は、1960年代に地域メッシュ統計の基盤としての地域メッシュ画定のための研究を省庁間共同プロジェクトとしてスタートさせ、1970年には「標準地域メッシュ体系」が告示された。その後、地域メッシュ統計については1973年に「統計に用いる標準地域メッシュおよび標準地域メッシュ・コード」が告示され、また1976年にはその作成に係る統一基準が「JIS C 6304」として策定された。このような動きと併行して、60年代後半には、センサスあるいは大規模標本調査に基づく地域メッシュ統計が整備された。メッシュによる統計表章の基盤となるのが、表1に示したような一連の区画体系を有する地域メッシュである。

表1中の第三次地域区画（基準地域メッシュ）からは、さらにその一辺を2等分、4等分、そして8等分した一辺約500mからなる1/2地域メッシュ、同じく250mの1/4地域メッシュが、さらに人口集中地区については、1/8地域メッシュ（同約125m）も作成されている。なお、これら各レベルの地域メッシュにはそれぞれ1桁のコードが付与されており、例えば最小単位である1/8地域メッシュは、全体で11桁からなるメッシュ・コードを持つ。

地域メッシュ統計は、電子地図、メッシュ枠情報、国勢調査区及び最小の地域単位である基本単位区の境界情報、それに事業所等の所在情報、さらには基本単位区内の調査票情報を用いて作成される。その作成は、事業所等の所在地又は基本単位区的位置を電子地図上に反映し、各個体の調査

票情報にそれらが所在するメッシュ・コードを対応させるという方法で行われる。なお、同定と呼ばれる調査票情報の各地域メッシュへの対応づけには、それらを地域メッシュに直接対応づける個別同定と調査区等の集計区域単位で対応づける調査区（基本単位区）同定とがある。

各メッシュに対応づけられた調査票情報の集合は、メッシュ・コード別に集計され、提供される。なお、一つの地域メッシュに表章される人口や世帯数が著しく少なく統計上の秘密が担保されないおそれがある場合には、人口総数など限定された集計結果だけが表章され、それ以外の諸項目については秘匿あるいは隣接するメッシュと統合の上で表示される。

社会統計の表現形態という側面から地域メッシュ統計を位置づけるとすれば、それは、当時の情報処理能力の制約の中で、統計データを空間情報とリンクさせることで空間的な分布情報を提供するものであるということになる。それは、基本的に行政区分に基づくそれまでの社会・人口統計指標にはない新たな形で分布情報の提供という意味で、社会人口統計体系(SSDS)の一種の外延的拡張であったといえる。

地域メッシュ統計は、メッシュという共通のプラットフォーム上で統計データと一部の行政情報とを接合することで防災地図や危険度マップ等の作成に活用されてきた。メッシュという物理的区画は、市町村合併のようにその時どきの行政区画の改変に左右されることなく、時系列比較あるいは空間的比較を可能にする。その一方で、地域メッシュについては、境界区分が行政区画に必ずしも対応していないことから、具体的な行政業務の遂行のための情報基盤としてはいま一つ適格的でないという利用面での制約を持っている。このため地域メッシュ統計は、統計の主要な表章形態として、その後必ずしも広範に普及したわけではない。

(3) 統計GIS

大量情報の処理を可能にする情報技術の発達は、メッシュからポリゴン(多角形)システムに基づく統計GISとして、小地域統計に新たな展開方向

を与えた。統計GISは、道路、鉄道、各種施設、土地利用区分などデジタル化された様々な空間データをレイヤー毎にデータベースとして維持、管理し、それを地図という共通のフィールド上で検索、加工処理することで、ある点や線からの一定範囲内の要素を抽出しバッファリング分析によって圏域人口やその特性等の算定を、地域メッシュ統計よりもはるかに高い精度で行うことを可能にした。現時点ではまだ利用可能な統計データの種類は限られているものの、統計GISは、地域メッシュ統計とは異なるポリゴンさらには曲線による空間的切り口による地域統計の表章を可能にするという点で、従来のSSDSのさらなる外延的展開として位置づけることができよう。

さらに統計GISは、地図という共通のプラットフォーム上で、統計と行政情報との接合をも可能にする。この点は、空間区画情報という制約こそあれ、集合値としての異種データ間の結合利用の際につきまとう種々の問題回避の糸口を与えるものでもある。

(4) 基本単位区

調査区は調査の正確でしかも円滑な実施のための調査技術上の仕組として設定されることから、通例、調査実施の都度その見直しが行われる。このような調査区の設定替えの結果、小地域統計の時系列的な比較可能性は担保されなくなる。そこでわが国では、平成2（1990）年国勢調査の調査区設定の際に、「小地域統計の時系列比較分析ばかりでなく、多様な地域区分による小地域統計の作成をも可能とする」との目的で、「調査結果の集計上の恒久的かつ最小の地域単位」〔(8) 5頁〕として基本単位区という地域区分が新たに導入された。

基本単位区は「固定的な境域として恒久化されて意味を持つ」〔(8) 5頁〕ことから、その区画は、住居表示に関する法律（昭和37年法律第119号）に基づく街区又は街区に準じた境域として設定され、市町村境界の変更、街区方式による住居表示の実施、町・字界の変更等による市町村内の住居区

画の変更、それに区画整理、地域再開発、水面埋立て、道路・河川等の新設・改修、災害等による地形・地物の著しい変化がない限り将来においても変更しないものとされている〔(8) 5頁〕。

このように、基本単位区は、時系列的比較可能性の要件を前面に出した調査結果の利用から出発した区域概念である点で、専ら調査技術上の区域概念である調査区とは明らかに一線を画している。しかし、他方で調査個票との関係で言えば、調査区よりもさらに小地域化されたものであるとはいえず、一個の基本単位区コードが複数の調査対象単位に関する調査票情報と対応しており、両者の間には依然として1対nの関係が維持されている。

4. 不完全な位置情報取得の統計的帰結—個体に帰属する位置情報と取得された位置情報との情報ギャップ

(1) data carrierとしての地域コードの特質

統計における調査票情報については、統計対象単位としての世帯や企業、事業所といった個体、すなわち個々の調査対象単位それ自体がいわゆるdata carrierとして、調査等で収集される様々な変数情報から構成されるdata bodyの担い手となっている点が、調査票情報に対して集計量とは全く異なる情報特性を付与する。すなわち、data bodyが氏名（名称）や住所（所在地）といった存在としての調査対象単位と一体化された個体識別情報によって担われているというまさにその点が、body情報に対して、識別情報あるいはそれと対応づけられた個人識別番号や企業コード等のID情報をキーとして個体単位での連結を可能にするという情報特性を新たに付与することになる〔(9)〕。

前節でも見てきたように、従来からの行政区による結果表章に加え、今日、多様な小地域集計結果が作成、提供されている。その集計結果をdata carrierとdata bodyという観点から捉えた場合、次のようになる。すなわち、各地域レベルについて様々な集計結果が提供されているが、それらは

当該地域レベルの地域コードをキーにそれと集計量の形でリンクされているのである。例えば、地域メッシュ統計の場合、各メッシュが画像情報におけるいわゆる画素に相当するdata carrierとして機能し、各メッシュに対応するメッシュ・コードをキーにdata bodyを構成する種々の統計情報が相互に関連づけられている。このことは、これらのコードをリンクキーとして該当する複数の個体に関するdata bodyを構成する変数群が相互に接合可能であることを意味する。ただし、data carrierの識別情報であるメッシュやポリゴン・コードと域内に所在する調査対象単位との間には一般に1対nの照応関係が成立していることから、data bodyの変数の接合は、個体ベースによるものではなく、あくまでも集計量としての接合利用可能性を保証しているだけである。

行政区、人口集中地区、調査区、基本単位区、それに各種メッシュやポリゴンによって区分される小地域には、それぞれの地域に対応するコードが付与されている。言い換えれば、これらの地域情報は、それぞれのコードをdata carrierとして持ち、相互にdata bodyがリンク可能なリレーショナルなデータ構造を持っている。ここでのデータ構造の特徴は、ひとつの地域コードに複数のdata bodyが同時にリンクされる1対nの対応関係にある点である。要するに、そこでは、地域コードというdata carrierが、data bodyとして複数の個体からなる集計値さらにはそれから導出された統計指標群を多次的にいわばレイヤーの形で背負っている。表2は、これらの関係を一覧表の形で整理したものである。

こういった様々な地域表章に共通するのは、該当区域内の調査対象単位の数の多寡こそあれ、複数の単位の位置情報が一個の地域コードによって統合して表示されている点である。data carrierをキー情報とする個体単位でのdata body情報の連結可能性という調査票情報の情報特性に照らして地域表章を捉えてみた場合、そこには調査区の設定という調査技術上の仕組みに制約された結果、調査票情報が本来保有すべき位置情報に対する明らかに不完全な情報の取得がある。なぜなら、そこでは、例えば、世帯、

事業所、あるいは企業といったそれぞれ特定の存在の場所を持つ個体情報が、基本的に調査区コード（あるいは基本単位区コード）として統合され、それに属する調査対象単位が一括して積み上げられることで各地域レベルの表章が行われているからである。

近年の小地域統計の整備の結果、地域別集計については、統計の解像度は飛躍的に改善した。しかし、「調査区」としての場所概念は、あくまでも地点を一定の面として捉えたものであり、調査地点情報について、統計対象単位が本来的に持つ統計原単位情報と調査個票に写し取られた調査票情報との間には明らかな情報の乖離が存在する。すなわち、統計の情報特性から見た場合、調査区の設定という実査のための調査技術上の仕組みが、位置情報についてその情報量を減衰させた形で調査票情報に投影させているのである。そこでは、統計対象単位が場所に関して本来的に保有する各地点情報は、空間（面）情報にアグリゲートされた形で記載されている。言い換えるなら、調査区コードは、地点情報に関して個体が持つ情報を喪失させた形で反映しているのである。

平成17年国勢調査の調査票には、調査員記入欄として、調査区内の世帯番号等とともに、4-1-2の合計7桁からなる調査区番号の記入欄が設けられている。調査項目に従業地又は通学地として、その所在地の都道府

表2 情報carrierとしての面情報と点情報（表）

ジオコード		データ キャリア	リレーシヨ ナルキー	属性・body変数 とのリンク形態	属性・body変数 とのリンク形式	body変数相互 の接合形態
空間	行政区	行政区コード	画素の空間 情報	1 対 n	地域別集計	集計量として の接合
	人口集中地区	DIDコード			小地域集計	
	調査区	調査区コード				
	基本単位区	基本単位区 コード			レイヤーによる 接合	
	メッシュ統計	メッシュ・コード				
	統計GIS	ポリゴン・コード				

県、市郡支庁、区町村記入欄はあるものの、常住地の所在地情報の記載欄はない。従って、国勢調査においては、調査個票のレベルで調査対象単位の所在を示す位置情報は調査区コードだけであり、統計調査の実査の際の調査員の独自記入情報も含め、収集された情報と実際の調査対象単位である世帯の位置情報との間には、復元不可能な情報喪失が存在する。

このような調査票情報における地点情報の喪失は、結果的に統計の利用可能性を大きく制約することになる。

(2) 問題の発生

基本単位区から調査区、町丁字、行政区分（市区町村、都道府県、地域ブロック、全国）へと地域集計の地域レベルの統合度が高くなるほど、平均化の作用により調査票情報が本来的に保有していた情報量は次第に減衰する。しかし、このような形で集計量が地域的に積み上げられる限り、集計に関して非整合的な問題は発生しない。しかし、世帯や事業所それに企業といった調査時点において住戸の住所や事業所や企業の所在地という地点情報（現実には微少な面情報）というdata carrierが本来的に有する位置情報としてではなく、調査区コードという n 個の要素が同一の位置コードを共有するという調査技術上の制約の下で把握された調査票情報における情報喪失は、次のような問題を内包する。

1) 行政区・調査区の再編

昭和や平成の大合併に象徴される市区町村の合併、あるいは市や政令指定都市への昇格のために隣接市町村の吸収合併が行われる場合、統計の時系列的比較可能性を担保するためには、新たな行政区分に従った遡及再集計が必要となる。その場合、より統合度の高い地域区分コードによる遡及は可能であっても、それが基本単位区の再編を伴う場合、統合された地域区分によって表章された集計結果のより低位の統合度を持つそれへの再分割は不可能である。地域が分割され複数の異なる地域に統合される場合には、新たな地域区分による過去の集計結果の遡及計算は大きな困難を伴う。

調査の実施に先立って調査区が設定される場合、道路や集合住宅等の構築物の新設などにより、しばしば調査区の再編、見直しが行われる。調査実施の技術的の必要に伴って発生する調査区の境界の変更は、調査区レベルでの時系列的比較可能性を担保しなくなる。恒久的性格を持つものとして導入された基本単位区についても、行政区の変更に対しては無関係ではあり得ない。

2) 利用面において発生する問題

上記の事例は、統計調査における調査客体の把握あるいは調査結果の積み上げに伴って発生する問題である。この他にも調査対象単位が本来的に持つ地点情報ではなくそれを統合した調査区コードとして位置情報が把握されていることがもたらす利用面での制約も存在する。それは、調査区による区分と異なる切り口での集計表章やデータの利用形態において顕在化する。

地域メッシュ統計の作成に際しては、メッシュというグリッド線によって調査区や基本単位区が複数の異なるグリッドに分割される場合、調査区というポリゴン内に位置する調査によって把握された調査票情報を各グリッドに配分するための同定作業が必要となる。また、統計GISにおけるバッファリング分析についても、それによるゾーニングと調査区のそれとが整合的ではない場合の方がむしろ一般的である。

5. GPS情報による調査票情報の情報価値の拡張可能性

(1) 統計利用面での拡張可能性

1) 時系列比較可能性の担保

すでに上述したように、情報処理能力の向上を受けて、例えば国勢調査の地域別表章においては、従来からの都道府県や市区町村といった行政区レベルでの域集計表の作成に加え、集計項目によっては人口集中地区さら

には町丁・字，調査単位区別の集計結果も作成，公表されてきた。

このような調査結果の地域別表章は，地域区分が変更された場合に時系列的な比較可能性が失われるという決定的な難点を持つ。旧地域区分による新たな調査結果の表章，あるいは新たな地域区分による過去の調査結果の遡及集計には多大の作業量を必要とすることから，国勢調査を例にとっても，それらは特別集計として一部の調査項目について行われているだけである。

行政区による地域設定，自然的あるいは人工構築物を境界とする地域設定につきまとうこの種の難点の回避策として導入されたのが，グリッドという方形によって物理的に地域の設定を行うメッシュ統計である。メッシュによる物理的空間設定を行うことによって，時系列的な比較可能性を確保するというのがその導入の契機であった。

地域メッシュ統計については，昭和48年の旧行政管理庁（現総務省）告示「統計に用いる標準地域メッシュおよび標準地域メッシュ・コード」（告示第143号）により，その統一的な作成方法が定められ，その後，日本工業規格「JIS C 6304」（その後「JIS X 0410」に改定）として制定された。

このようなメッシュによる地域区分の規格化により，メッシュ・コードは，当該小地域に関する調査結果，行政情報，さらには各種推計をレイヤーとしてあるいは集計量として相互にリンクすることのできるdata carrierとしての性格を獲得した。これによって，水害や地震被害に対するハザードマップ等の作成や高齢化率，外国人居住者率といった関係比率による小地域の特徴の視覚的表示を可能にするという点で，メッシュ統計はそれまでの集計表とは異なる異種の情報源情報の連結利用可能性を提供することになった。

その反面でメッシュ統計は以下のような難点を持つ。その1は，地域区分についての自由度の欠如である。国勢調査から作成されるメッシュ統計については，平成17年調査で人口集中地区について五次メッシュが作成されるようになっているが，五次メッシュが作成されている地域についても，

最大で5つのレベルの地域区分が利用できるだけである。

第2の問題は、事業所などの所在地又は世帯の調査区（基本単位区）の調査対象単位の各地域メッシュへの対応づけという同定作業の必要性である。なお、事業所データについてはすでに個々の事業所の所在地情報を日本測地系の住所辞書データから緯度、経度情報を取得してメッシュコードを付与するアドレスマッチングによる同定（「アドレス同定」）が行われている。

しかしながら、国勢調査の地域メッシュ統計の作成にあたっては、これまで包含同定、面積割合同定、中心点同定（図心同定、人口分布点同定）といった方法により、メッシュの境界をまたいで存在する基本単位区については、その人口の各グリッドへの配分が行われてきた。

世帯に関する調査票情報についても事業所の場合と同様に、各住戸の住所に基づきGPSコードを付与できるとすれば、以下のような分析が可能になる。まず、仮に、行政区画・調査区の境界が変更されても、任意時点での行政区画に従って時系列方向での串刺しの製表ができる。もしGPSコードを付与した形で世帯に関する調査票情報をアーカイビングすることができれば、より弾力的な形で時系列的な地域表章が可能になる。その場合、仮に時間の経過の中で行政区分、調査区の境界が変更されたとしても、任意のゾーンング（ポリゴン）による時間軸での串刺しの比較領域の設定が可能になる。さらには、地域的に統合された調査票情報の集団を改めて再分割することも可能になる。そこでは、任意の時点での地域区分による時系列的集団構成が可能となり、遡及集計も機械的に行うことができる。

GPSコードが付与された異なる複数の時点の世帯や事業所等の調査票情報をGPSコードによりリンクすることによって構成されるデータ系列は、もちろん厳密な意味での縦断面データではない。それは、いわば擬似的なパネルデータである。しかしそれらに企業・事業所名、世帯主名などの情報を追加的に使用することで、企業動態や世帯動態の分析データとして用いることができる。

2) GPS付レコードの追加的な利用可能性

調査票情報にGPSコードを付与することによって、利用者は、従来の調査票情報が持っていた調査区コード⁽¹⁾や基本単位区コード⁽²⁾に起因する様々なデータの利用制約から解放されることになる。事業所や企業の所在地住所情報の測位システムによる緯度、経度情報へのコンバートによる位置確定は、わが国でもすでに事業所企業統計調査データによる地域メッシュ統計の作成の際の各メッシュへの同定作業において用いられている。このように、事業所・企業に関する調査票情報が持つ所在地情報をGPSコード化することによって、既存の調査票情報が持つ情報の様々な活用可能性を拡張することができる。

(i) 事業所・企業に関する調査票情報

(a) クロスセクションデータによる分析の可能性

ある特定時点のクロスセクション型のGPS付の調査票情報を用いることによって、業種別の産業立地・集積状況、その規模（生産額、従業員規模）、集積密度等を集積地域別に把握することができる。また、それらを時系列的に比較することで、集積度等の時間的変化を追跡することができる。

(b) パネル型データによる分析の可能性

わが国の場合、事業所や企業を調査対象単位とする統計調査の多くでは、これまでそれぞれ独自の事業所コード等のID番号が付与されてきた。このため、個々の調査レベルでは個体レコード間の縦断的リンケージが部分的に可能であった。しかしながら、わが国では、事業所の把握はあくまでも調査区ベースで行われてきた。その結果、企業や事業所の新規参入と転入移転、あるいは廃業や倒産による退出と転出移転の区別ができないという問題を抱えている。

この種のパネルデータに各調査時点における調査対象単位の所在地を示すGPS情報を持たせることで、ID番号をマッチングキーとして異時点のレコードをマッチングすることにより、企業や事業所の継続、新規参入、退出状況を明確に把握することができるようになる。

その結果、このタイプのデータを用いることによって、次のような分析も可能となる。すなわち、同一の企業、事業所が異時点で異なるGPS情報を持っていれば、調査対象単位の空間的移動を意味する。このような空間的移動状況を業種別、規模別等に分析することで、企業や事業所の立地の再編あるいは集積状況の動態的変化状況をつかむことができる。さらに、これらの移動に関するOD行列を作成することで、移動の方向における特徴を把握することもできよう。

(c) その他の利用可能性

例えば、事業所ベースの調査票情報を被災地域と電子地図と対応させることで、手作業による照合作業を行うことなく被害推計が可能である〔(10)〕。

(ii) 世帯に関する調査票情報

GPSコードの付与による調査票情報の利用可能性の拡張は、事業所の場合と同様に、理論的には世帯に関する調査票情報に対しても適用可能である。それは、ハザードマップ等と連動させた被害予測、被害推計といった防災上の利用可能性だけでなく、いわゆるバッファリング分析による小地域マーケティングへの利用可能性を拓くものである。そこでは、地域メッシュ統計の場合と異なり、任意の地点を中心とする任意の地域形状による地域の切り取り（バッファリング）に対応しうることで、その利用可能性は飛躍的に拡充される。

(2) 統計作成面

1) 所在地の特定による実査の精度向上

調査員の目視によって企業や事業所といった調査対象単位の把握を行ってきた企業・事業所統計調査では、経済のサービス化やICT化の進展に伴い、事業形態の多様化やSOHO等の増加により調査漏れとなるケースが拡大してきている。これについては、税務情報や企業登記から得られる経済活動を営む事業所等の所在地情報が利用できれば、それを測位系情報に変

換することで、具体的な所在地の特定を行うことができる。これによって、単なる調査員の目視による把握漏れに起因する回収率の低下部分についても調査面での対応が可能となる。また、これによってビジネス・レジスターに格納される情報の維持更新に必要なプロファイリングの対象となる事業所や企業の特特定もより容易となる。

一方、世帯調査については、例えばアメリカにおける人口センサスの実施部局である商務省センサス局では、2010年人口センサス改善計画（2010 Population Census Enhancement Program）に基づき、センサス実施の約1年前にあたる2009年4月から7月にかけてArcPad付ハンドヘルドPCを携帯した調査員が各住戸を訪問し、住戸および集合施設のGPS情報を取得するaddress canvassingを実施した。これによって取得したGPS情報をバーコードとして貼付した調査票が、MAF（Master Address File）の住所情報を用いて各戸に郵送された。調査期日後一定期間を経ても調査票が郵送提出されていない住戸についてはTIGER電子地図上に表示され、調査員がそのピンポイント表示情報を参照しつつ調査票の回収に当たった。このようなGPSの利用は、長期低落傾向にある人口センサスの回収率の回復につながるものとして期待されている。

2) マッチングキーとしての使用

フランスのビジネス・レジスターSIRENEのように、各企業や事業所に対して統一的なコードが付与され、調査や行政記録へのその使用が義務づけられている場合、調査の結果得られた調査票情報や行政記録のリンケージは容易である。しかし、このような制度整備ができていない国はなお少数にとどまり、多くの国では複数のID体系が並存している。

このような国では、名称、所在地、電話番号等の情報から照合確率を導出して行う統計的マッチングという方法が一般に採用されている。その場合、より高い照合確率をもつ組合せの順に照合作業が行われ、機械的照合ができなかったものについては、最終的に手作業による照合（clerical matching）が行われる。しかしながら、件数が膨大な場合、照合の相手を

数多くの候補の中から選択する必要がある、照合作業は、極めて労働集約的作業となる。統一ID番号が整備されていない場合、あるいは照合可能なID番号リストが不在の場合、個体レコード相互のリンケージにはこのような作業が不可欠となる。

そこで、所在地あるいは住所情報をGPSコードにコンバートすることができれば、GPSコードを有力な照合キー情報として使用でき、あるいは仮に照合確率が低位であっても、照合相手を比較的狭い空間に存在する対象群に限定することで、手作業による照合の作業効率を大幅に改善できるものと期待される。その意味で、GPSコードは、マッチングの精度あるいはその作業効率の向上にとって有効であると考えられる。特に日本の場合、現状では①統一企業コードの未整備、②税務情報（VAT）が使用できないという制度上の制約がある中で、ビジネス・レジスターを構築するためには、調査データと名簿情報とのマッチングによる情報のリンケージが不可欠である。アーカイブに格納された個々の調査票情報が位置情報を有する場合、GPSコードあるいは秘匿性の観点でそれから導出した代理変数をリンクキー変数として用いることで、調査データ同士あるいは調査データと行政データのリンケージによる新たな統計作成が可能となる。

3) 地点情報を軸とするパネルデータの作成可能性

GPSコードをマッチングキーとして相互に接合された同一系列の調査データあるいは行政情報は、例えば、同一の所在地で経済活動を営む企業（事業所）、あるいは同じ住戸に居住する世帯に関する時間的経過の中での統計的記録を与えることができる。しかしながら、GPSを用いてリンクされた複数のレコードはあくまでも地点情報を共有しているだけで、直ちに特定の個体についての縦断面データを与えるものではない。その意味で、このようにして編成されたデータは、一種の擬似パネルデータである。

同じGPSコードを持つ複数の同じ変数群から構成される調査票情報について、統計対象単位である世帯や事業所等が同じ名称あるいは氏名、ID番号等の識別情報を持つ場合、それらは同一個体の経年的変化を表すデータ

であるといえる。そこには、継続的に営業活動を営んでいる活動主体あるいはライフステージを更新した世帯の姿が投影されている。このように、GPSコード情報を用いてリンクされた擬似パネルデータに他の個体識別情報を適用することで、そこから純正の縦断面データを編成することができる。

4) business demography指標作成への利用可能性

上記3)の擬似パネルデータセットの中には、リンクされた調査票情報の識別情報の同一性が確認できないケースが存在しうる。特に企業や事業所に関する情報の場合、それらは、当該活動拠点におけるある活動主体の退出と新たな主体の参入といった時点間での個体の交代を示唆している。例えば、産業（業種）、業態、立地別にそれらを比較分析することにより、現在OECDやEurostatなどが中心的に推進しているbusiness demography分析に対して新たな視点からの指標構築の可能性を提案することができる。

5) 人口移動の統計的把握への利用可能性

世帯に関して調査等から得られた調査票情報が把握時点における居住地のGPSコード情報とともにいわゆるGPS付きの調査票情報としてデータベースに格納されている場合、世帯の識別情報によって各レコードが持つGPSコード情報を名寄せすることによって、その世帯の空間的移動に関する情報を得ることができる。

6) レイヤーの差し替えによる拡張された調査票情報編成の可能性

GISは、デジタル地図という共通のプラットフォーム上にレイヤーとして様々な集計量、非集計量を重ねて表示したり、また各レイヤーが持つ情報を相互に関連づけることで新たなレイヤーを編成することができる。同様のことは、GPSという地点（正確に言えば、微細な面）情報を共通のプラットフォームとして用いることで、同じGPS情報を持つ多様な情報源の情報を多次元的に接合することができる。これによって、情報源を異にする調査票情報が持つ変数は、レコード編成上、統合され拡張された調査票情報の各変数とみなすことができ、それらを交叉的に併用した分析が可能

になる。これによって、個々に独立したいいわゆるstand alone型の統計の場合には果たせなかった共変量要素をコントロールした統計分析なども可能となる。

むすび

本稿では、調査対象としての企業、事業所そして世帯といった調査対象単位が本来保有する空間位置情報に焦点をあてることで、個票調査の特徴並びにそれから得られる統計原単位情報が本来的に持つ情報価値の拡張可能性について考察してきた。さいごに、今回の考察から得られた論点の整理を行うとともに、今後に残された課題を指摘することで、本稿でのむすびとしたい。

本文で既に述べたように、表式調査においては、表式への記入が行われる統計の把握時点ですでに集計量として取り扱われていることから、個体に帰属する統計原単位情報は単なる集計量の一構成要素に過ぎず、独立した存在とはなっていない。このように、調査対象単位が個体そのものとして調査実施者の意識に上っていないことから、当然のことながら、それが社会的存在である個体として本来的にまとっている場所情報についても、直接的な意識の対象とはなりえない。表式調査の場合、報告が上申され、府県や国といったより上位の地域レベルの表章において、その内訳として個々の報告結果が地域を代表する統計となった時点においてはじめて場所情報は意味を持つのである。

一方、個票調査の場合、マイヤーの立論に代表されるように、現実の調査過程では、調査対象単位を正確に1回しかも過不足なく把握するために、調査が実施される空間的範囲を排他的でしかも空白部分のない形で調査区に区分し、その把握が行われてきた。わが国の統計史に詳しい鯨島は、個票調査の意義について、母集団把握との関係で次のように指摘している。「終戦期までのわが国の統計思想の中には、母集団の概念が根本的に欠けて

いた。より一般的に言えば、集団概念についての認識が欠けていたのだ。集団とそれを構成する単位概念が官庁統計の中に認識されてきたのは、ようやく大正期にはいつてからのことである。つまり、日本の統計思想の一般水準は、大正期にはいつて、ようやく点計調査の概念、すなわち集団を構成する単位を点計することによって集団の量、特性を知るという認識段階に達したのであった」〔(3) p.279〕、と。

しかし、ここで鮫島が「単位を点計する」というときの点計概念には、単に調査対象単位を個体に帰属する統計原単位情報の痕跡を留めつつ統計に編成することを意味しているだけで、そこには時間とともに個体の存在を規定する位置に関する「点計」概念は全く含意されていない点に留意する必要がある。鮫島においても調査対象単位はその存在のポイントを意味する位置情報としてではなく、依然として調査区コードによって表現される小空間として「面計」されている。世帯、事業所それに企業といった調査対象単位は、一般に住戸や建物等の構築物という存在の形態を持ち、それらはいわば調査区という面の内部にその存在の場所を持つ点（あるいは微小空間）と対応づけることができる。それにもかかわらず、個票調査が導入されて以来今日まで、位置情報は、調査対象単位そのものが本来的に有する地点情報としてではなく、調査区（あるいは基本単位区）として、調査票を介してその把握が行われてきた。

調査区内に位置する複数の調査対象単位が共通の調査区（あるいは基本単位区）コードをその位置情報として持つという調査技術上の要件は、その必然的帰結として、客体が本来的に保有する統計原単位情報と調査票に写し取られたそれとの間に、取得された位置情報に関する情報損失を発生させることになる。調査過程におけるこの情報損失が様々な形で調査結果の利用制約要因を形作っていることはすでに本文で指摘したとおりである。

統計作成過程にGPS技術を導入することにより、調査対象単位を文字通り「点計」することが可能となる。これは個票調査がそもそも内在的に保有していた調査論理であり、GPSという最新の情報技術がはじめてそれを

現実的に可能にしたものである。

GPS技術を統計の作成・利用面に適用することによって、調査票情報の潜在的利用可能性の多方面での開拓が可能となるように思われる。本稿では、その方法論的検討を主たる課題としたことから、GPS技術の具体的な利用方法、GPS情報をレイヤーとする既存の情報の多様な拡張をはじめとして、その多様かつ具体的な利用可能性の検討については、今後の課題としたい。

〔注〕

- (1) 調査区コードは、調査区を識別する主番号（最大4桁）、調査区の種類を示す後置番号（1桁）、それに基本単位区を識別するための単位番号（最大2桁）からなりそれらをハイフンでつないだ最大7桁の数値置情報から構成される。〔(8) 4頁〕
- (2) 基本単位区は4桁からなる町字コードと5桁の基本単位区コードとから構成される。このうち前者は、町あるいは大字単位の一連番号を示す3桁と分割番号を示す1桁とからなる。一方、基本単位区コードの冒頭の2桁は丁目や小字などに対応する一連番号であり、次の2桁が基本単位区の一連番号、最後の1桁は分割番号を意味している。〔(8) 6頁〕

〔参考文献〕

- (1) 蛭川虎三『統計利用に於ける基本問題』岩波書店 1932年
- (2) 蛭川虎三『統計概論』岩波書店 1934年
- (3) 相原茂、鯉島龍行『統計日本経済』筑摩書房 1971年
- (4) 上杉正一郎『経済学と統計（改訂新版）』青木書店 1974年
- (5) 木村太郎『統計・統計方法・統計学』産業統計研究社 1977年
- (6) 森博美『統計法規と統計体系』法政大学出版局 1991年
- (7) 森博美「我が国政府統計の展開と展望」『統計』日本統計協会 2007年1月
- (8) 総務省統計局『平成17年国勢調査 調査区関係資料 利用の手引き』2007年
- (9) 森博美「統計個票情報の情報特性について」『経済志林』法政大学経済学部学会 第76巻第4号 2009年
- (10) Georg von Mayr (1914), *Statistik und Gesellschaftslehre*, (大橋隆憲訳『統計学の本質と方法』)
- (11) Jarmin S. Ron and Miranda J., (2009) “The Impact of Hurricanes Katrina, Rita and Wilma on Business Establishments: A GIS Approach” *Journal of Business Valuation and Economic Loss Analysis*. Vol.4

本論文は、平成22年度日本学術振興会科学研究費補助金 基盤B「政府統計データのアーカイビングシステムの構造と機能に関する国際比較研

究」(課題番号 22330070研究代表者 森博美)による研究成果の一部である。

Exploring New Dimensions of Statistical Information by Furnishing GPS Codes on Individual Records

Hiromi MORI

《Abstract》

Statistical practices in countries such as U.S. and France address that geographical information has been widely accepted not only as statistical GIS but also in the arena of producing the qualified data. The present author has attempted, under the input raised by a set of latest developments in statistical practices, the methodological review on the possible expansion of the usability of individual records armed by the GPS coordinates.

Apart from table-based surveys which were major way of collecting statistical information at the dawn of the history of modern statistical survey, statistical surveys in modern area have been characterized by questionnaires returned from respective respondents. In the questionnaire-based surveys, however, census tract codes identified the surveyed units such as households, establishments and enterprises in terms of geographical information not as individual entity but as groups, which a number of survey units share the identical tract code.

Standing on the basic idea that returns in surveys represent statistical information inherent in each respondents, the present author has tried in this paper to explore new dimensions of usability by replacing tract codes with GPS codes that correspond to each individual survey unit. The proposed new statistical coding methodology has not only evidenced several restrictions in use of survey results derived from traditional spatial coding system but also put under light many possibilities of exploring potentials of statistical data.